60 Rec'd PCT/PTO 24 SEP 2001

FORM PTO-1390 (REV 5-93)	U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEY DOCKET NO. 100564-00080
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		DATE: September 24, 2001
		U.S. APPLN. NO. (IF KNOWN, SEE 37 C.F.R. 1.5) Not Yet Assigned 09/926201
INTERNATIONAL APPLICATION NO. PCT/EP00/02607	INTERNATIONAL FILING DATE 23 March 2000	PRIORITY DATE CLAIMED 23 March 1999 and 18 October 1999

TITLE OF INVENTION: DETECTION, CLONING AND SEQUENCING OF POLYPEPTIDES WHICH DRIVE THE SUBCELLULAR LOCALIZATION OF PROTEINS

APPLICANT(S) FOR DO/EO/US: Caystano GONZALEZ, Luis BEJARANO

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371. (THE BASIC FILING FEE IS ATTACHED)
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☒ This express request to begin national examination procedures [35 U.S.C. 371(f)] at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☐ A proper demand for International Preliminary Amendment was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed [35 U.S.C. 371(c)(2)]
   a. ☒ is transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☐ has been transmitted by the International Bureau.
   c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☒ A translation of the International Application into English [35 U.S.C. 371(c)(2)].
7. ☒ Amendments to the claims of the International Application under PCT Article 19 [35 U.S.C. 371(c)(3)]
   a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☐ have been transmitted by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ A translation of the amendments to the claims under PCT Article 19 [35 U.S.C. 371(c)(3)].
9. ☐ An oath or declaration of the inventor(s) [35 U.S.C. 371(c)(4)].
10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 [35 U.S.C. 371(c)(5)].

Items 11 - 16 below concern other document(s) or information included:

11. ☒ An Information Disclosure Statement under 37 C.F.R. 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 C.F.R. 3.28 and 3.31 is included.
13. ☒ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information: ☒ PCT/ISA/210 International Search Report; PCT/RO/101 Pct Request; PCT/IPEA/409 International Preliminary Examination Report; 1 sheet of Amended Claims; International Application as Published WO 00/56875 Drawings (8 sheets)

U.S. APPLN. NO. (IF KNOWN SEE 37 C.F.R. 1.50) Not Yet Assigned 09/926201	INTERNATIONAL APPLICATION NO. PCT/EP00/02607	ATTORNEY DOCKET NO. 100564-00080 DATE: September 24, 2001

17. ☒ The following fees are submitted:

Basic National Fee [37 C.F.R. 1.492(a)(1)-(5)]:

Search Report has been prepared by the EPO or JPO.....$860.00
International preliminary examination fee paid to USPTO (37 C.F.R. 1.482).....$690.00
No international preliminary examination fee paid to USPTO (37 C.F.R. 1.482) but international search fee paid to USPTO [37 C.F.R. 1.445(a)(2)].....$710.00
Neither international preliminary examination fee (37 C.F.R. 1.482) or international search fee [37 C.F.R. 1.445(a)(2)] paid to USPTO.....$1,000.00
International preliminary examination fee paid to USPTO (37 C.F.R. 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4).....$ 100.00

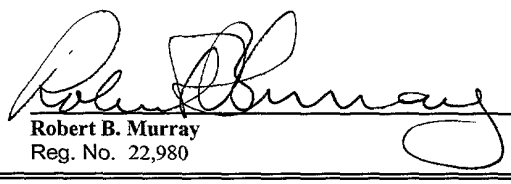
	CALCULATIONS	PTO USE ONLY
ENTER APPROPRIATE BASIC FEE AMOUNT =	$ 860.00	
Surcharge of $130.00 for furnishing the oath or declaration later than ☐ 20 ☒ 30 months from the earliest claimed priority date [37 C.F.R. 1.492(e)].	$ 130.00	

Claims	Number Filed	Number Extra	Rate		
Total Claims	22 - 20 =	2	X $ 18.00	$ 36.00	
Independent Claims	1 - 3 =	0	X $ 80.00	$ 0.00	
Multiple dependent claim(s) (if applicable)			+ $270.00	$ 0.00	
TOTAL OF ABOVE CALCULATIONS =				$ 1,026.00	
Reduction by one-half for filing by small entity, if applicable. Verified Small Entity statement must also be filed. (Note 37 C.F.R. 1.9, 1.27, 1.28).			-	$ 0.00	
SUBTOTAL =				$ 1,026.00	
Processing fee of $130.00 for furnishing the English translation later the ☐ 20 ☐ 30 months from the earliest claimed priority date [37 C.F.R. 1.492(f)].			+	$ 0.00	
TOTAL NATIONAL FEE =				$ 1,026.00	
Fee for recording the enclosed assignment [37 C.F.R. 1.21(h)]. The assignment must be accompanied by an appropriate cover sheet (37 C.F.R. 3.28, 3.31). $40.00 per property			+	$ 0.00	
TOTAL FEES ENCLOSED =				$ 1,026.00	
				Amount to be refunded	$
				Charged	$

a. ☒ A check in the amount of $1,026.00 to cover the above fees is enclosed.
b. ☐ Please charge my Deposit Account No. 01-2300 in the amount of $ to cover the above fee. A duplicate copy of this sheet is enclosed.
c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 01-2300.

NOTE: Where an appropriate time limit under 37 C.F.R. 1.494 or 1.495 has not been met, a petition to revive [37 C.F.R. 1.137(a) or (b)] must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:
Arent Fox Kintner Plotkin & Kahn
1050 Connecticut Avenue, N.W.
Suite 400
Washington, D.C. 20036-5339
Tel: (202) 857-6000 Fax: (202) 638-4810
RBM/aam



Robert B. Murray
Reg. No. 22,980

JC09 Rec'd PCT/PTO 24 SEP 2001
09/926201

PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of:
C. GONZALEZ et al.

Appln. No.: PCT/EP00/02607

Filed: Concurrently herewith Attorney Dkt. No.: 100564-00080

For: DETECTION, CLONING AND SEQUENCING OF POLYPEPTIDES WHICH DRIVE THE SUBCELLULAR LOCALIZATION OF PROTEINS

PRELIMINARY AMENDMENT

Commissioner for Patents
Washington, D.C. 20231 September 24, 2001

Sir:

Prior to calculation of the filing fees and initial examination of the application, please amend the above-identified application as follows:

IN THE CLAIMS:

Please amend claims 3, 4, 5, 7, 9, 10, 12, 15, 18, 20, 21, and 22 as follows:

3. (Amended) Process according to claim 1

characterized in that a cDNA or cDNA fragments are used as random nucleic acids.

4. (Amended) Process according to claim 1,

characterized in that a homologous system of library and cells for the transfection is used.

5. (Amended) Process according to claim 1,

characterized in that a heterologous system of library and cells for the transfection is used.

7. (Amended) Process according to claim 6,

characterized in that a reporter gene leading to a visually detectable signal upon expression is used.

9. (Amended) Process according to claim 1,

characterized in that the vector contains an inducible promoter driving the expression of random nucleic acid and marker gene.

10. (Amended) Process for the identification and/or production of a protein that is localized in a given subcellular localization,

characterized in that a nucleic acid coding for a polypeptide or part thereof driving the localization in said given subcellular localization is cloned according to claim 1 and the nucleic acid is used to detect DNA sequences coding for a protein containing such polypeptide or part thereof.

12. (Amended) Process for directing the subcellular localization of a nucleic acid expression product,

characterized in that a polypeptide driving the localization of a protein containing such polypeptide or part thereof is detected, its nucleic acid sequence is obtained by a process according to claim 1, the nucleic acid coding for the polypeptide or part thereof is fused to a nucleic acid coding for a protein to be expressed, and the fusion product is expressed.

15. (Amended) Process according to claim 12,

characterized in that the fusion product contains a proteolytic cleavage site between the protein to be expressed and the polypeptide or part thereof and/or reporter gene product.

16. (Amended) Use of a polypeptide or part thereof, which drives the subcellular localization of a protein containing such polypeptide or part thereof, and which is detected and/or cloned according to claim 1 in a vector for the expression of a desired protein wherein the vector contains a specific site into which a DNA encoding said desired protein can be inserted,

characterized in that the vector further comprises a DNA sequence encoding a polypeptide or a part thereof which drives the subcellular localization of a protein containing such polypeptide or part thereof, which DNA sequence is positioned in such a way that a fusion protein of desired protein and polypeptide or part thereof is encoded.

18. (Amended) Vector according to claim 16,

characterized in that the vector further comprises a reporter gene positioned in such a way that a fusion protein of desired protein and polypeptide or part thereof and reporter gene product is encoded.

20. (Amended) Vector according to claim 16,

characterized in that the vector further contains sequences encoding proteolytic cleavage sites between one or more of the constituents of the fusion protein.

21. (Amended) Cell line,

characterized in that it is transfected with a vector according to claim 16, encoding a fusion protein of at least a polypeptide or part thereof driving the localization to a given subcellular localization and a desired protein.

22. (Amended) Kit for the expression of a desired protein in a desired localization of a host cell,

characterized in that it contains a vector according to claim 16 optionally together with other components and/or buffers for the protein expression.

Please add the following claim:

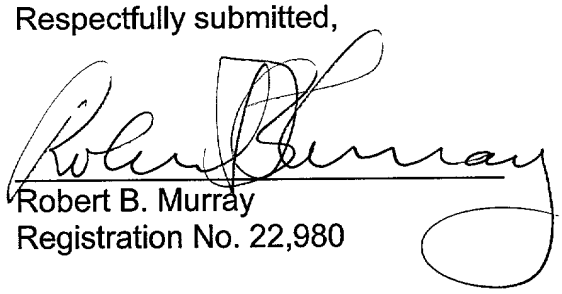
--24. Kit for the expression of a desired protein in a desired localization of a host cell,

characterized in that it contains a cell line according to claim 21 optionally together with other components and/or buffers for the protein expression.--

REMARKS

Claims 1-23 are pending in this application. By this Amendment, claims 3, 4, 5, 7, 9, 10, 12, 15, 18, 20, 21, and 22 are amended to correct the multiple dependency thereof and claim 24 has been added to place this application into better condition for examination. No new matter is added.

Respectfully submitted,



Robert B. Murray
Registration No. 22,980

ARENT FOX KINTNER PLOTKIN & KAHN, PLLC
1050 Connecticut Avenue, N.W.,
Suite 400
Washington, D.C. 20036-5339
Tel: (202) 857-6000
Fax: (202) 638-4810
RBM:aam/epb

Marked Up Claims

JC12 Rec'd PCT/PTO 24 SEP 2001

09/926201

Claims

1. Process for the detection, cloning and/or sequencing of polypeptides or parts thereof, which drive the subcellular localization of a protein containing such polypeptide or part thereof,
characterized in that the process comprises the following steps:
   (a) constructing an expression library of random nucleic acids ligated to a reporter gene and contained in a vector molecule,
   (b) transfecting a plurality of host cells with the library,
   (c) screening for the subcellular localization of the expression product of the nucleic acid in the host cells via detection of a signal produced by the reporter gene,
   (d) cloning such cells where the reporter gene signal is detected in a certain subcellular localization, and
   (e) cloning and optionally sequencing the nucleic acid insert which encodes the polypeptide or part thereof.

2. Process according to claim 1,
characterized in that a cDNA or cDNA fragments are used as random nucleic acids.

3. Process according to claim 1 [or 2]
characterized in that a eukaryotic or a yeast library is used.

4. Process according to [anyone of claims 1 to 3] claim 1
characterized in that a homologous system of library and cells for the transfection is used.

5. Process according to [anyone of claims 1 to 3] claim 1
characterized in that a heterologous system of library and cells for the transfection is used.

6. Process according to claim 5,
characterized in that a Drosophila library is used to transfect mammalian or yeast cells.

7. Process according to [anyone of claims 1 to 6] claim 1
characterized in that a reporter gene leading to a visually detectable signal upon expression is used.

8. Process according to claim 7,
characterized in that nucleic acids coding for GFP, BFP, luciferase or YFP are used as reporter gene.

9. Process according to [anyone of claims 1 to 8] claim 1
characterized in that the vector contains an inducible promoter driving the expression of random nucleic acid and marker gene.

10. Process for the identification and/or production of a protein that is localized in a given subcellular localization,
characterized in that a nucleic acid coding for a polypeptide or part thereof driving the localization in said given subcellular localization is cloned according to [claims 1 to 9] claim 1 and the nucleic acid is used to detect DNA sequences coding for a protein containing such polypeptide or part thereof.

11. Process according to claim 10,
characterized in that for the production of the protein the nucleic acid is expressed in an expression system.

12. Process for directing the subcellular localization of a nucleic acid expression product,
characterized in that a polypeptide driving the localization of a protein containing such polypeptide or part thereof is detected, its nucleic

acid sequence is obtained by a process according to [anyone of claims 1 to 8] the nucleic acid coding for the polypeptide or part thereof is fused to a nucleic acid coding for a protein to be expressed, and the fusion product is expressed. claim 1

13. Process according to claim 12,
characterized in that a nucleic acid coding for the polypeptide or part thereof and a reporter gene is fused to the nucleic acid coding for a protein to be expressed.

14. Process according to claim 12,
characterized in that a reporter gene the expression product of which is visually detectable is used.

15. Process according to [anyone of claims 12 to 14] claim 12
characterized in that the fusion product contains a proteolytic cleavage site between the protein to be expressed and the polypeptide or part thereof and/or reporter gene product.

16. Vector for the expression of a desired protein wherein the vector contains a specific site into which a DNA encoding said desired protein can be inserted,
characterized in that the vector further comprises a DNA sequence encoding a polypeptide or a part thereof which drives the subcellular localization of a protein containing such polypeptide or part thereof, which DNA sequence is positioned in such a way that a fusion protein of desired protein and polypeptide or part thereof is encoded.

17. Vector according to claim 16,
characterized in that the vector is a eucaryotic vector.

18. Vector according to claim 16 [or 17]

characterized in that the vector further comprises a reporter gene positioned in such a way that a fusion protein of desired protein and polypeptide or part thereof and reporter gene product is encoded.

19. Vector according to claim 18,
characterized in that the reporter gene product is visually detectable.

20. Vector according to [anyoner of claims 16 to 19] claim 16
characterized in that the vector further contains sequences encoding proteolytic cleavage sites between one or more of the constituents of the fusion protein.

21. Cell line,
characterized in that it is transfected with a vector according to claim 16 [anyone of claims 16 to 20] encoding a fusion protein of at least a polypeptide or part thereof driving the localisation to a given subcellular localisation and a desired protein.

22. Kit for the expression of a desired protein in a desired localisation of a host cell,
characterized in that it contains a vector according to [anyone of claims 16 to 20] claim 16 or a cell line according to claim 21 optionally together with other components and/or buffers for the protein expression.

23. Collection of cell lines according to claim 21.

13 Rec'd PCT/PTO 21 DEC 2001
09/926201

PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

GONZALEZ et al. Atty. Dck. No. 100564-00080

Serial No.: 09/926,201

Filed: September 24, 2001

For: DETECTION, CLONING AND SEQUENCING OF POLYPEPTIDES WHICH DRIVE THE SUBCELLULAR LOCALIZATION OF PROTEINS

STATEMENT UNDER 37 C.F.R. §1.821

Commissioner for Patents
Washington, D.C. 20231 December 21, 2001

Sir:

In accordance with 37 CFR §1.821, applicants hereby submit the Sequence Listing for the above-referenced application in paper copy and computer readable form.

The name of the file on the computer readable for is 100564-00080.txt. the paper copy and computer readable form are the same, and no new matter has been added.

In the vent that this paper is not considered timely filed, applicants hereby petition for an appropriate extension of time. If necessary, please charge any additional amounts or credit any overpayments to Deposit Account No. 01-2300.

Respectfully submitted,

Lynn A. Bristol
Registration No. 48,898

Customer No. 004372: ARENT FOX KINTNER PLOTKIN & KAHN, PLLC
1050 Connecticut Avenue, N.W., Suite 400
Washington, D.C. 20036-5339
Tel: (202) 857-6000; Fax: (202) 638-4810
LAB/ccd

SEQUENCE LISTING

<110> Gonzalez, Cayetano
Bejarano, Luis

<120> DETECTION, CLONING AND SEQUENCING OF POLYPEPTIDES WHICH DRIVE THE SUBCELLULAR LOCALIZATION OF PROTEINS

<130> 100564-00080

<140> 09/926,201
<141> 2000-03-23

<150> PCT/EP00/02607
<151> 2000-03-23

<150> EP 99105846.2
<151> 1999-03-23

<150> EP99120622.8
<151> 1999-10-18

<160> 8

<170> Patent-In version 3.1

<210> 1
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(26)
<223> Description of Artificial Sequence: oligonucleotide (Oligo A)

<400> 1
catgttggcg gccgcggtac cgtcga 26

<210> 2
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(23)
<223> Description of Artificial Sequence: oligonucleotide (Oligo B)

<400> 2
gcccgggcgt gagcaagggc gag 23

<210> 3
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(16)
<223> Description of Artificial Sequence: oligonucleotide (Oligo Not 1-1b)

<400> 3
gatcgcggcc gcgtac 16

<210> 4
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(8)
<223> Description of Artificial Sequence: oligonucleotide (Oligo Not 1-8)

<400> 4
gcggccgc 8

<210> 5
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(27)
<223> Description of Artificial Sequence: Fir

<400> 5
agcttcgaat tcgcggccgc caacatg 27

<210> 6
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(27)
<223> Description of Artificial Sequence: Sec

<400> 6
tatgatctag agtcgcggcc gctttac 27

<210> 7
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(27)
<223> Description of Artificial Sequence: Thi

<400> 7
tagcgctacc ggactcagat ctcgagc 27

<210> 8
<211> 27

<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(27)
<223> Description of Artificial Sequence: Fou

<400> 8
aaaacctcta caaatgtggt atggctg 27

Date Recue/Date Received 2020-05-26

B/PRTS

JC09 Rec'd PCT/PTO 24 SEP 2001

09/926201

Detection, cloning and sequencing of polypeptides which drive the subcellular localization of proteins

Specification

The present invention relates to a process for the detection, cloning and/or sequencing of polypeptides or parts thereof, which drive the subcellular localization of a protein containing such polypeptide or part thereof, a process for the identification and/or production of a protein that is localized in a given subcellular localization, and a process for directing the subcellular localization of a nucleic acid expression product.

One of the most conspicuous features of the eukaryotic cell is its high degree of compartmentalization. Chromatin, nuclear matrix, nuclear membrane, Golgi apparatus, endoplasmatic reticulum, the endo- and exocytic compartments, the actin and microtubule cytoskeletons, mitochondria, the centrosome and the cell membrane are just some of the major subcellular organelles/compartments which have been defined by standard cytological analysis. Moreover, most of these compartments can be further subdivided into well differentiated regions or structures having a cytological and molecular identity of their own, thus resulting in the large number of subcellular domains which characterize the eukaryotic cell.

The physiological relevance of such compartmentalization is paramount. Every major cellular activity can be assigned to one or more well defined subcellular compartments. As a matter of fact, the intricate regulatory networks which operate within eukaryotic cells greatly rely on the differential subcellular localization of the molecules involved. The close relationship between subcellular localization and function is such that in most instances determining the preferential subcellular localization of a protein provides one of the best clues as to its putative function.

The molecular basis of specific subcellular localizations is not yet well understood. In some cases the localized protein contains a functional domain which drives its targeting either directly or through interaction with other previously localized members of the target structure. Well known examples of this first case are the nuclear localization signals which are recognized by the nuclear pore complex and translocated into the nucleus or the combination of a polybasic domain and a C-terminal CAAX motif, which leads to the post-translational modification of the protein and its membrane targeting. The second case includes kinesins and MAPs the cytoskeletal, spindle or centrosomal localization of which is achieved by virtue of their interaction with microtubuli. Finally, it is also conceivable that other mechanisms, e.g. different rates of import, export and degradation, could result in steady-states which may account for the preferential subcellular localization of proteins which do not contain any bona fide subcellular localization signal of their own.

In view of the diversity of mechanisms which may account for the subcellular localization of a given protein, it was an object of the present invention to provide a possibility to detect signals which either directly or indirectly drive subcellular localization. Such signals are usually polypeptides or parts thereof that are present in proteins. The knowledge of such polypeptides can be useful for a plurality of assays or applications.

The object underlying the invention was accomplished by the provided process for the detection, cloning and/or sequencing of polypeptides or parts thereof, which drive the subcellular localization of a protein containing such polypeptide, wherein said process comprises:

(a) constructing an expression library of random nucleic acids ligated to a reporter gene and contained in a vector molecule,

(b) transfecting a plurality of host cells with the library,

(c) screening for the subcellular localization of the expression product of the nucleic acid in the host cells via detection of a signal produced by the reporter gene,

(d) cloning such cells where the reporter gene signal is detected in a certain subcellular localization of interest, and

(e) cloning and optionally sequencing the nucleic acid insert which encodes the polypeptide or part thereof.

The above process allows for the detection of polypeptides or parts thereof, which drive the targeting of a protein to a particular subcellular location or structure, said process being completely independent of the function, organization and length of the respective protein containing such polypeptide. The process according to the invention can also be used to detect polypeptides or parts thereof that relocate intracellulary under specific conditions, like stress for example following heat shock or infection with a pathogen, all of which result in a dramatic rearrangement of the architecture of the cell (Cudmore et al., Trends Microbiol. (1997) 5, 142-147). Finally also polypeptides or parts thereof can be detected or cleaved according to the method of the invention that mediate the retention of proteins at specific organelle structures or loci.

Depending on the length of the random nucleic acids the process of the invention also allows the detection of complete or nearly complete proteins, that due to the presence of a polypeptide driving the localization are transfered to a certain intracellular location after expression. According to the invention random nucleic acids are produced from the genome of an organism. Either genomic DNA or cDNA may be used to generate the random nucleic acids that are ligated to a reporter molecule and inserted in vector molecules to construct the expression library of feature (a). Random nucleic acid molecules that are produced by subjecting the DNA of interest to restriction cleavage are only detected in the location of interest if they contain at least such a portion of said polypeptide which ensures that the

subcellular localization driving function is retained. Other constructs might also lead to expression of a fusion protein displaying the reporter gene signal albeit not in the localization of interest.

Out-of-frame insertions are generally not expressed in the process according to the invention.

It is irrelevant whether the random nucleic acid and reporter gene are ligated before being inserted in a vector molecule together, or whether a vector containing the reporter gene (lateron also termed GET (GFP-epitope trap) vector) is constructed into which a random nucleic acid can easily be ligated in an appropriate location next to the reporter gene. It is generally preferred to produce a fusion gene product, with the reporter gene located at the C-terminus. If the reporter gene is located at the N-terminal end, it will always be expressed thereby increasing background. Nevertheless, this might be the only way to detect some proteins which will not get localized if the fusion is made in the other direction. Theoretically, it is also possible that the reporter gene product is located in between nucleic acid expression products, as long as both the reporter gene and/or the polypeptide or part thereof retain their function.

The thus obtained vector molecules containing the random nucleic acid and reporter gene are transfected into a plurality of host cells which are then subjected to conditions allowing for the expression of the vector insert, whereupon screening for the subcellular localization of the expression product of the nucleic acid in the host cells takes place via detection of a signal produced by the reporter gene. Cells which show the reporter gene signal in a subcellular localization of interest are selected and subcloned, whereupon the DNA sequence insert can be cloned and optionally sequenced, said nucleic acid insert encoding the polypeptide or part thereof.

A schematic presentation of the process according to the invention is shown in Fig. 1. Fig. 1 shows the GFP-epitope trap approach. Random DNA fragments cloned into the GET vector will produce fusion proteins between the polypeptides encoded by these inserts and GFP. The GET vector contains a high efficiency cloning site immediate after the initial ATG of the GFP which shifts this codon out of frame with the rest of the coding sequence. Thus, GFP can only be expressed from vectors carrying an insert that restores the reading frame. Upon transfection with the GET library, a fraction of the cells can be observed to express GFP which in some cases will be localised in a particular compartment or organelle. These cells can be cloned and the inserts that code for the relevant subcellular localisation signal can be isolated by RT-PCR.

In a preferred embodiment of the present invention, a cDNA is used to create the random nucleic acids, or an expression library made of a cDNA is used.

In a further preferred embodiment of the invention, a library, either genomic or cDNA, from a mammalian organism or yeast or C. elegans or C. laevis is used. This preferred embodiment is not intended to limit the invention, since DNA from any organism may be used to detect specific polypeptide or parts thereof which drive the subcellular localization in the respective organism.

In one embodiment of the invention, a homologous system of nucleic acid for the creation of the expression library and host cell for the transfection is used. The homologous expression system is meant to identify a system where host cells are used that belong to the same species from which the nucleic acid was obtained.

In another embodiment of the invention, a heterologous system of nucleic acid library and cells for the transfection is used. In such a case, the host

cell does not belong to the same species as the nucleic acid that is to be expressed therein.

An example for a heterologous system would be the use of a Drosophila DNA for the generation of the expression library and of mammalian or yeast cells as host cells to be transfected with the vector molecule.

For the process according to the invention, standard procedures can be used which are known to the man in the art. Cloning of cells can be done either manually, picking up the cells and replating them for as many times as required to isolate one clone, or by serial dilution. Also, a fax sorter may be used which separates individual cells expressing a specific marker.

In a preferred embodiment of the invention, a reporter gene leading to a visually detectable signal upon expression is used.

Although principally other reporter genes are suitable, too, the visually detectable expression product is most easily detectable. Especially preferred reporter genes are genes coding for GFP (green fluorescent protein), or GFP derivatives like for example BFP (blue fluorescent protein), luciferase, YFP (yellow fluorescent protein), or CFP (cyan fluorescent protein). These derivatives are described in Pepperkok et al., Current Biology 9: 269 - 272 (1999) and references quoted therein.

The process according to the invention makes it possible to establish a system allowing to screen a huge number of nucleic acid molecules for the presence of a sequence encoding a polypeptide or part thereof capable of driving the subcellular localization of a protein containing such polypeptide. The process according to the invention has also the advantage that it can preferably be used in higher eukaryotes.

Therefore, a further subject of the present invention is a process for identifying and/or producing a protein that is localized in a given subcellular location, such process comprising the above process according to the invention as well as the cloning of a nucleic acid coding for a polypeptide epitope driving the localization in this given subcellular localization, and the use of said cloned nucleic acid to detect longer DNA sequences coding for a protein containing such polypeptide epitope. This detection can be conducted by standard molecular biology techniques, for example by hybridizing the nucleic acid to a genomic or cDNA library from a certain species and detecting homologous sequences encoding a protein, by RT-PCR or by comparing the obtained nucleic acid sequence with databases containing a huge number of DNA sequences. Such databases contain sequences coding for known proteins as well as sequences which are postulated to be coding for a protein, which, however, has either not been identified yet or the function of which is still unknown. The process according to the invention, therefore, is also useful as a high throughput method of determination of the subcellular localisation of the fast growing number of sequences which are being generated or detected by ongoing genome projects.

As soon as, by the above process for the identification of a protein containing the polypeptide, the respective nucleic acid coding for such protein is obtained or the sequence thereof is known, said nucleic acid can be expressed in an expression system, producing said protein containing a polypeptide or used in any other way including formation of mutants etc.

Another application of the process according to the invention is the identification of the proteins that are differentially sorted in differentiating cells, like for example cells that are induced to polarize or primary cultures of differentiating neurons (Dotti and Simmons, Cell (1990), 62: 63-72).

A further interesting use of the present process will also be the cloning of interacting partners of a given protein by transfecting cells which contain the protein labelled with a fluorochrome that produces FRET (Cubitt, et al., TIBS (1995) 20: 448-455) with the library's reporter.

Finally, systematic screenings with a library produced according to the invention could be used to identify new domains within known organelles and compartments.

A still further subject of the present invention is a process for directing the subcellular localization of a nucleic acid expression product. Said process comprises detecting a polypeptide or part thereof driving the localization of a protein containing such polypeptide according to the above process of the invention, and obtaining the nucleotide sequence encoding such polypeptide or part thereof, wherein further the nucleic acid coding for the polypeptide or part thereof is fused to a nucleic acid coding for a protein to be expressed, and the fusion product is expressed.

In a preferred embodiment of the present invention, the nucleic acid for the polypeptide or part thereof driving the localization and a reporter gene are fused with the nucleic acid coding for a protein to be expressed. By such fusion of a polypeptide and a reporter gene with a nucleic acid, the actual expression of the protein to be expressed at the localization of interest can be monitored.

For this purpose, it is preferred that a reporter gene is used the expression product of which is visually detectable.

It is a further preferred but in no way obligatory embodiment of the invention that the fusion product of the protein to be expressed and polypeptide or part thereof and/or reporter gene contains a proteolytic cleavage site between the protein to be expressed and the polypeptide

and/or the reporter gene product. According to this preferred embodiment, it is possible to obtain the pure protein to be expressed in a given localization by cleaving off the part directing the localization and optionally also the part enabling the monitoring of the expression and localization.

To this end, it might be useful to also express a corresponding proteolytic enzyme and direct it to the same subcellular localization by means of the process of the invention.

A further subject matter of the present invention is a vector for the expression of a desired protein wherein the vector contains a specific site into which a DNA encoding said desired protein can be inserted, said vector being characterized by further comprising a DNA sequence encoding a polypeptide or a part thereof which drives the subcellular localisation of a protein containing such polypeptide or part thereof, which DNA sequence is positioned in such a way that a fusion protein of desired protein and polypeptide or part thereof is encoded.

According to the invention any vector which is suitable for gene expression in an envisaged expression system can be employed. In a preferred embodiment of the invention, the vector is a eucaryotic vector and the envisaged expression system a eucaryotic system. The specific site into which a DNA encoding said desire protein can be inserted preferably is a restriction site that allows an in frame expression of the DNAs encoding the desired protein and encoding the polypeptide or part thereof. The site can also be a polylinker containing several restriction sites.

As described above for the process of directing the subcellular localization of a nucleic acid expression product, also for the vector according to the invention it is preferable that the vector further contains a reporter gene in such a manner, that upon expression of the desired protein and the polypeptide driving the localization or the part thereof also a reporter gene

product is expressed. This reporter gene product can either be expressed in form of a fusion protein with the other two components, or it can be separately expressed as a separate fusion with polypeptide driving the localization or part thereof.

In all of these vector constructs encoding fusion proteins of polypeptide driving the localization and/or reporter gene it is further preferable that DNA sequences encoding proteolytic cleavage sites are present is such a position that after expression the components can be separated from each other, thus facilitating purification of the desired protein. In this connection a further vector may contain a gene coding for the proteolytic enzyme also in such a manner that it is connected with a polypeptide driving the intracellular localization or part thereof. It is also feasable that the same vector coding for a fusion protein also contains the DNA sequences necessary to encode a fusion which results in a proteolytic enzyme being expressed and localized to the same cellular compartment as the other fusion protein, containing the desired protein.

It is also possible to use the process according to the invention for the detection, cloning and/or sequencing of polypeptides or parts thereof, which drive the subcellular localization of a protein containing such polypeptide or part thereof, for establishing a cell line or a collection of cell lines which are tranformed with a vector according to the invention. Such cell lines may show a reporter gene product at different locations. Such cell lines or such collection of cell lines is a further subject of the present invention as well as a kit containing a vector or a cell line according to the invention and which is useful for the expression of a desired protein in a desired localisation of a host cell.

The following examples along with the accompanying figures are intended to further elucidate the invention:

Fig. 1 shows a presentation of the general process of the invention.

Fig. 2 A - D shows a schematic representation of the process steps performed in Example 1.

Fig. 3 shows examples of subcellular localization of a reporter protein.

Fig. 4 shows patterns of GFP localisation generated by transfection with a GET library. Low (A) and high (B to I) magnification views of HEK 293 cells were counterstained for DNA using propidium iodine (red). B) mitochromosomes (arrows); F) the mitotic spindle. We have not determined yet the subcellular localisation of GFP in the cells shown in panels 2 G, H and I. Scale bar = 15 μm.

Fig. 5 shows sequencing the inserts that target GFP localisation. A) The GFP fusion in clone 02/11#22 shows a strong nucleolar localisation with a faint homogeneous nuclear background. B) The insert from this clone contains a well defined bipartite NLS (red) and meets the consensus of a nucleolar localisation signal. C) In clone 09/07#18 GFP colocalised with the ER as shown by counterstanding with an antibody against α-calnexin (not shown). D) The insert from this cell line encodes a peptide of 35 amino acids that contains a predicted trans-membrane motif (PMSIFQLIYFLLFLFLGVIC). This sequence does not have a match in the sequence databases. Scale bar = 15 μm.

Example 1

Creating the Td2 fragment in EGFP-N1

The vector pEGFP-N1 (CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303-4230, USA) was modified by PCR using the following primers:

Oligo A (5'-CATGTTGGCGGCCGCGGTACCGTCGA-3')* (SEQ.ID.NO.1)

Oligo B (5'-GCCCGGGCGTGAGCAAGGGCGAG-3'). (SEQ.ID.NO.2)

Oligo A contains an ATG with a good Kozak and a SrfI site. The ATG is out of frame with the GFP coding sequence.

PCR was carried out with Expand High Fidelity PCR System (Boehringer Mannheim GmbH, Sandhofer Strasse 116, D-68305 Mannheim, Germany) as indicated in the following protocol:

Step	Temperature	Time
1	94°C	3 min
2	42°C	1 min
3	68°C	4 min (slope 13°C/sg)
4	94°C	1 min
5	42°C	1 min
6	68°C	4 min (goes to 4, 7 cycles)
7	68°C	7 min
8	4°C	pause

The PCR product was purified using the PCR Quiaquick Purification System (QIAGEN GmbH, Max-Valmer-Strasse 4, 40724 Hilden, Germany) and ligated with Rapid Ligation Kit (Boehringer Mannheim GmbH, supra). The ligated vector was used to transfect Epicurian Coli XL1-Blue (Stratagene, 11011 North Torrey Pines Road, La Jolla, CA 92097) by heat shock. After transfection, the cells were plated out in LB-Agar supplemented with 30 μg/ml kanamycin and incubated for 16 hours at 37°C. The DNA of some of the resulting colonies was isolated by minipreps and analyzed with restriction enzymes to confirm that it corresponded to the expected modified vector.

From this modified vector, the Td2 fragment was purified and subcloned into the NotI site of vectors of the pQE30 series (previously modified to contain this site). As expected, only Td2 fragments cloned into the pQE30 NotI resulted in cells expressing GFP. After this check, the Td2 fragment was cloned back into the NotI site of pEGFP-N1NotI. This is a derivative of pEGFP-N1 into which a NotI site was introduced in the polylinker. The final vector is pEGFPTd2.

Modification of the pQE30 series

These were modified introducing between the Bam HI and KpnI sites an adapter containing a NotI site. The adapter was made by annealing the following oligos:

Not 1-1b (5'-GATCGCGGCCGCGTAC-3') (SEQ.ID.NO.3)

Not 1-8 (5'-GCGGCCGC-3'). (SEQ.ID.NO.4)

Fig. 2 A - D shows the procedures of Example 1 schematically.

Example 2

Construction of the "epitope-trap" library

Drosophila gDNA was cut to completion with AluI and HaeIII (Boehringer Mannheim, supra), purified with QIAEXII Gel Extraction Kit, run in agarose gel for further size selection, purified again with QIAEXII Gel Extraction Kit and cloned into the SrfI of pEGFPTd2. Ligated DNA was used to transform E. coli XL1-Blue MR (Stratagene, supra). A small fraction of the cells was plated in LB medium containing 30 μg/ml kanamycin. The resulting clones were isolated and their DNA was purified and analyzed to determine the size of the average insert. The results were about 420,000 clones, approx. 6,700 (1.6%) of which had no insert, the estimated average length of the inserts being around 490 bp.

The remaining cells were plated out over a sterile nylon filter laid on a 24x24 plate containing LB supplemented with kanamycin, incubated at 30°C for 24 hours, replicated into another filter and reincubated for 4 hours at 37°C. The DNA was then obtained using Plasmid Maxi kit by QIAGEN.

Cells were transfected with a library prepared as described above. Ten hours after transfection, cells were fixed with methanol, and observed under the microscope. The reporter protein used in this example (GFP) can be observed as a bright white. In each of the examples shown in Fig. 2, the reporter is localized in different components within the cells. Observations were made with a Leica TCS confocal microscope system (Leica, Germany).

Example 3

In a typical transfection experiment with HEK293 cells and the MmcDNA-GET library, about 50% of the cells express GFP of which 20% display a distinct localisation of this reporter. Around 8 to ten hours after transfection, some cells start to express GFP and the first localisation patterns are recognisable (Figure 4A). Figure 4B to I shows some of the GFP localisation patterns that we observed. Panel 4B shows GFP specficially localised in the mitochondria as confirmed by counterstaining with the mitochondria-specific marker mitotracker (not shown). In the cell shown in Figure 4C GFP displays is fairly uniform in the cytoplasm, but is significantly concentrated in a small area near the nucleus that corresponds to the centrosome (arrow) as revealed by counterstaining with a human, autoimmune anti-centrosome antibody (not shown). Panels 4D, E and F show mitotic cells from different GFP expressing lines. GFP can be seen to localise in the cytokinesis furrow (arrow; Figure 4D), the chromosomes (Figure 4E) and the mitotic spindle (Figure 4F). GFP does not appear to be localised during interphase in these two cells. We have not yet determined the precise subcellular localisation of GFP in the cells shown in Panels 4G, H and I.

These are a few examples of the patterns of GFP localisation that we have observed. Using GET we have been able to identify cells with GFP localised in every major organelle and compartment. These observations illustrate the power of GET to identify specific molecular associations to organelles and compartments. To demonstrate the use of GET to identify proteins sequences that carry targeting signals we have cloned and sequenced the DNA inserts from some of these cells. As expected, we have found sequences that correspond to known proteins and contain targeting signals which are consistent with the observed localisation of the GFP fusion. One of these is clone 02/11#22, (Figure 5A, B). The GFP fusion in this cell line shows a distinct nucleolar localisation with a weak nuclear background. The insert from this line is identical to a fragment that spans between amino acids 62 and 131 of the mouse homologue of the HTLV-I tax responsive element binding protein TAXREB107 (Nacken et al., Biochim Biophys Acta (1995), 1261: 432-434). This fragment contains a well defined bipartite nuclear localisation signal (KRKYSAAKTKVEKKKKKE) and meets the consensus of a nucleolus localisation signal. We have also found inserts that are new sequences which do not have a match in the databases. This is the case of clone 09/07#18 (Figure 5C, D). These cell contain GFP that is tightly localised to the endoplasmic riticulum (ER), as shown by counterstaining with an antibody against the ER marker α-calnexin (not shown) (Cannon et al., J. Biol. Chem. (1999), 274: 7537-7544). The insert from this cell line encodes a peptide, 35 amino acids long. It does not have a match in the sequence databases, but contains a predicted trans-membrane motif (PMSIFIQLIYFLLFLFLGVIC) that may occur for the ER specific retention shown by the fusion protein (Dotti et al., Cell (1990), 62: 63-72.

Example 4

Construction of the GET vector (GET#1). Using primers A (C A T G T T G G C G G C C G C G G T A C C G T C G A) and B

(GCCCGGGCGTGAGCAAGGGCGAG) we modified pEGFP-N1 (Clontech) by PCR to introduce a SrfI site between nucleotides three and four of the GFP coding sequence. This insertion shifts the initial ATG codon of the GFP out of frame with the rest of the coding sequence. This ensures that only insert-containing plasmids will express GFP. PCR was carried out with the Expand High Fidelity PCR System (Boehringer). Oligo A also introduced a Not1 site 10 nucleotides upstream of the GFP CDS. The PCR product was purified using the PCR Quiaquick Purification System (Quiagen), ligated with Rapid Ligation Kit (Boehringer) and used to transform Epicurian Coli XL1-Blue (Stratagene), by heat-shock. Transformed cells were plated out in LB-Agar supplemented with 30 μg/ml kanamycin and incubated for 16 hours at 37°C. The modified vector was then isolated by minipreps and the Not1 fragment subcloned into a pQE31 vector previously modified to introduce a Not1 site between the BamHI and KpnI sites with an adaptor made with oligos Not1-1b (GATCGCGGCCGCGTAC) and Not1-8 (GCGGCCGC). The resulting colonies were checked under a transiluminator to test the expression of GFP and the Not1 fragment was then isolated from one of the colonies and subcloned into pEGFP-N1-Not, a modified version of pEGFP-N1 that carries an additional Not1 site inserted in position 635-642.

Example 5

Construction of the MmcDNA-GET library. The cDNA was obtained from NIH/3T3 cells by random priming, purified with QIAEX II Gel Extraction Kit, cloned into the SrfI site of the GET#1 vector using the Rapid Ligation Kit (Boehringer) and transformed into E. coli XL1-Blue MR (Stratagene). Plating out a small aliquote of these cells we estimated that the library contained about 420.000 clones of which 1.6% had no insert. The complete library was then plated out onto a sterile Nylon filter laid out on a 24x24 cm plate containing LB supplemented with kanamycin, incubated at 30°C for 24 hours, replicated into another filter and reincubated for 4 hours at 37°C.

The DNA from the library was then purified with Plasmid Maxi Kit (QIAGEN).

Example 6

Transfection of HEK293 cells with GET libraries and cloning of cells displaying localised GFP. HEK293 cells were transfected with the GET libraries following the method described by Chen and Okayama (Mol Cell Biol (1987), 7: 2745-2752). The cells were observed twelve to sixteen hours after transfection to check for localised GFP using an inverted LEICA DMI-RBE microscope using a long distance 63x Fluotar objective. The position of the cells of interest was labelled with a diamond pen and then cloned by a combination of manual cloning and serial dilutions, as described in Harlow and Lane (Antibodies: a laboratory manual (1988), Cold Spring Harbour Laboratory Press, N.Y.). In some cases, the cells were first cloned using a fluorescence-activated cell sorter (FACS) and the resulting clones were later analysed to determine the presence of localised GFP.

Example 7

Cloning of the DNA fragments encoding subcellular localisation sequences. These were isolated from cloned cells by RT-nested PCR using oligos Fir (AGCTTCGAATTCGCGGCCGCCAACATG) Sec (TATGATCTAGAGTCGCGGCCGCTTTAC) Thi (TAGCGCTACCGGACTCAGATCTCGAGC) and Fou (AAAACCTCTACAAATGTGGTATGGCTG) which flank the SrfI site of the GET#1 vector. mRNA isolation was carried out using the mRNA Capture Kit (Boehringer). The reverse transcriptase reaction and the first round of PCR were carried out using the Titan One Tube RT-PCR Kit with the Expand High

Fidelity PCR System (Boehringer). Oligo Fou was used to prime the RT reaction. The first and second rounds of PCR used oligos Thi and Fou and Fir and Sec as primers. The PCR product was run on an agarose gel and isolated with QIAEX II Gel Extraction Kit (QIAGEN). The isolated fragment was then digested with Not 1 and subcloned into the GET#1 vector to check that the isolated fragment drives the GFP to the expected localisation and for sequencing.

Claims

1. Process for the detection, cloning and/or sequencing of polypeptides or parts thereof, which drive the subcellular localization of a protein containing such polypeptide or part thereof,
characterized in that the process comprises the following steps:
(a) constructing an expression library of random nucleic acids ligated to a reporter gene and contained in a vector molecule,
(b) transfecting a plurality of host cells with the library,
(c) screening for the subcellular localization of the expression product of the nucleic acid in the host cells via detection of a signal produced by the reporter gene,
(d) cloning such cells where the reporter gene signal is detected in a certain subcellular localization, and
(e) cloning and optionally sequencing the nucleic acid insert which encodes the polypeptide or part thereof.

2. Process according to claim 1,
characterized in that a cDNA or cDNA fragments are used as random nucleic acids.

3. Process according to claim 1 or 2,
characterized in that a eukaryotic or a yeast library is used.

4. Process according to anyone of claims 1 to 3,
characterized in that a homologous system of library and cells for the transfection is used.

5. Process according to anyone of claims 1 to 3,
characterized in that a heterologous system of library and cells for the transfection is used.

6. Process according to claim 5,
characterized in that a Drosophila library is used to transfect mammalian or yeast cells.

7. Process according to anyone of claims 1 to 6,
characterized in that a reporter gene leading to a visually detectable signal upon expression is used.

8. Process according to claim 7,
characterized in that nucleic acids coding for GFP, BFP, luciferase or YFP are used as reporter gene.

9. Process according to anyone of claims 1 to 8,
characterized in that the vector contains an inducible promoter driving the expression of random nucleic acid and marker gene.

10. Process for the identification and/or production of a protein that is localized in a given subcellular localization,
characterized in that a nucleic acid coding for a polypeptide or part thereof driving the localization in said given subcellular localization is cloned according to claims 1 to 9 and the nucleic acid is used to detect DNA sequences coding for a protein containing such polypeptide or part thereof.

11. Process according to claim 10,
characterized in that for the production of the protein the nucleic acid is expressed in an expression system.

12. Process for directing the subcellular localization of a nucleic acid expression product,
characterized in that a polypeptide driving the localization of a protein containing such polypeptide or part thereof is detected, its nucleic

acid sequence is obtained by a process according to anyone of claims 1 to 8, the nucleic acid coding for the polypeptide or part thereof is fused to a nucleic acid coding for a protein to be expressed, and the fusion product is expressed.

13. Process according to claim 12,
characterized in that a nucleic acid coding for the polypeptide or part thereof and a reporter gene is fused to the nucleic acid coding for a protein to be expressed.

14. Process according to claim 12,
characterized in that a reporter gene the expression product of which is visually detectable is used.

15. Process according to anyone of claims 12 to 14,
characterized in that the fusion product contains a proteolytic cleavage site between the protein to be expressed and the polypeptide or part thereof and/or reporter gene product.

16. Vector for the expression of a desired protein wherein the vector contains a specific site into which a DNA encoding said desired protein can be inserted,
characterized in that the vector further comprises a DNA sequence encoding a polypeptide or a part thereof which drives the subcellular localization of a protein containing such polypeptide or part thereof, which DNA sequence is positioned in such a way that a fusion protein of desired protein and polypeptide or part thereof is encoded.

17. Vector according to claim 16,
characterized in that the vector is a eucaryotic vector.

18. Vector according to claim 16 or 17,

characterized in that the vector further comprises a reporter gene positioned in such a way that a fusion protein of desired protein and polypeptide or part thereof and reporter gene product is encoded.

19. Vector according to claim 18,
characterized in that the reporter gene product is visually detectable.

20. Vector according to anyoner of claims 16 to 19,
characterized in that the vector further contains sequences encoding proteolytic cleavage sites between one or more of the constituents of the fusion protein.

21. Cell line,
characterized in that it is transfected with a vector according to anyone of claims 16 to 20, encoding a fusion protein of at least a polypeptide or part thereof driving the localisation to a given subcellular localisation and a desired protein.

22. Kit for the expression of a desired protein in a desired localisation of a host cell,
characterized in that it contains a vector according to anyone of claims 16 to 20 or a cell line according to claim 21 optionally together with other components and/or buffers for the protein expression.

23. Collection of cell lines according to claim 21.

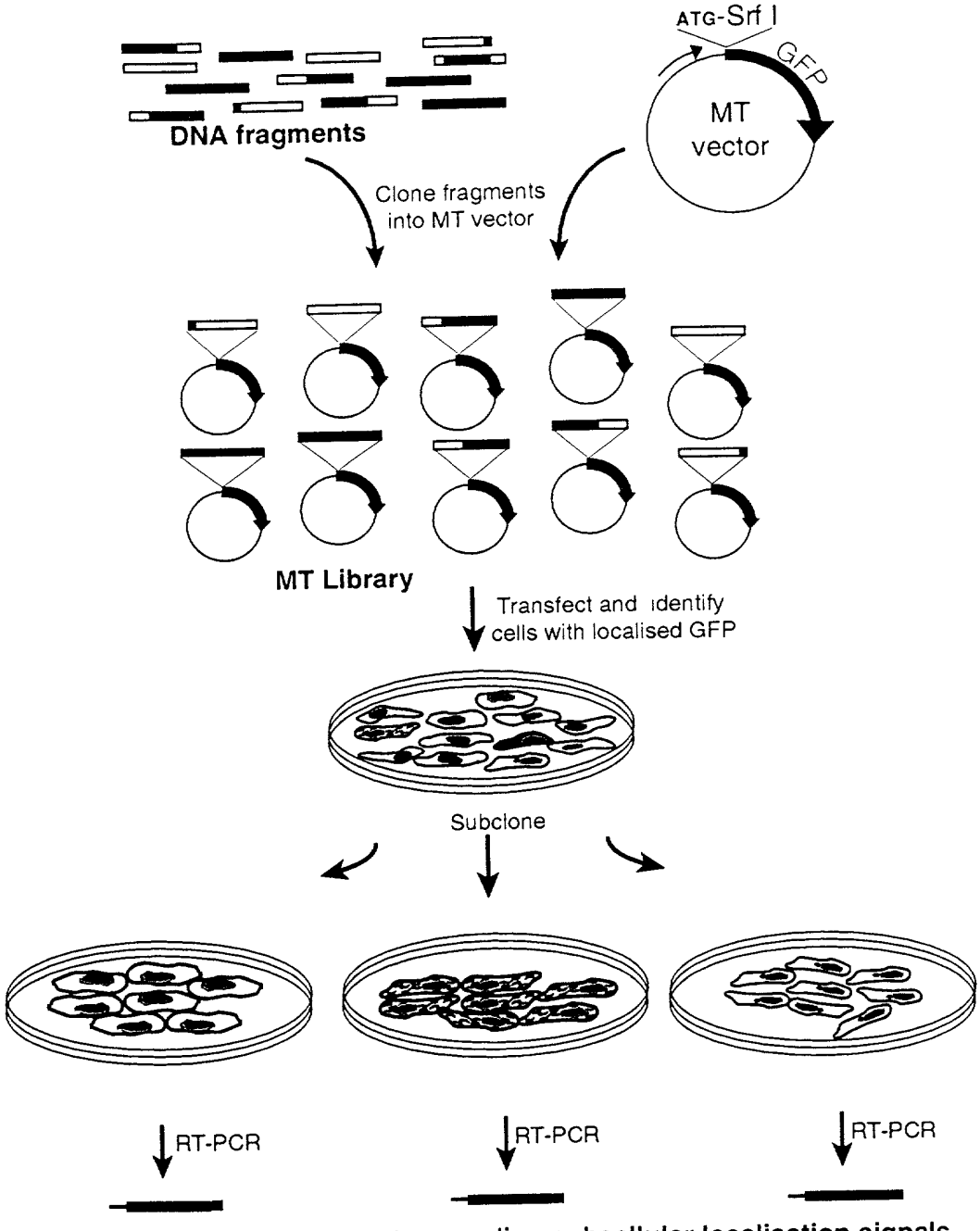

ATG-Srf I
GFP
MT
vector
DNA fragments
Clone fragments
into MT vector
MT Library
Transfect and identify
cells with localised GFP
Subclone
RT-PCR
RT-PCR
RT-PCR
Isolated DNA fragments encoding subcellular localisation signals

Figure 1

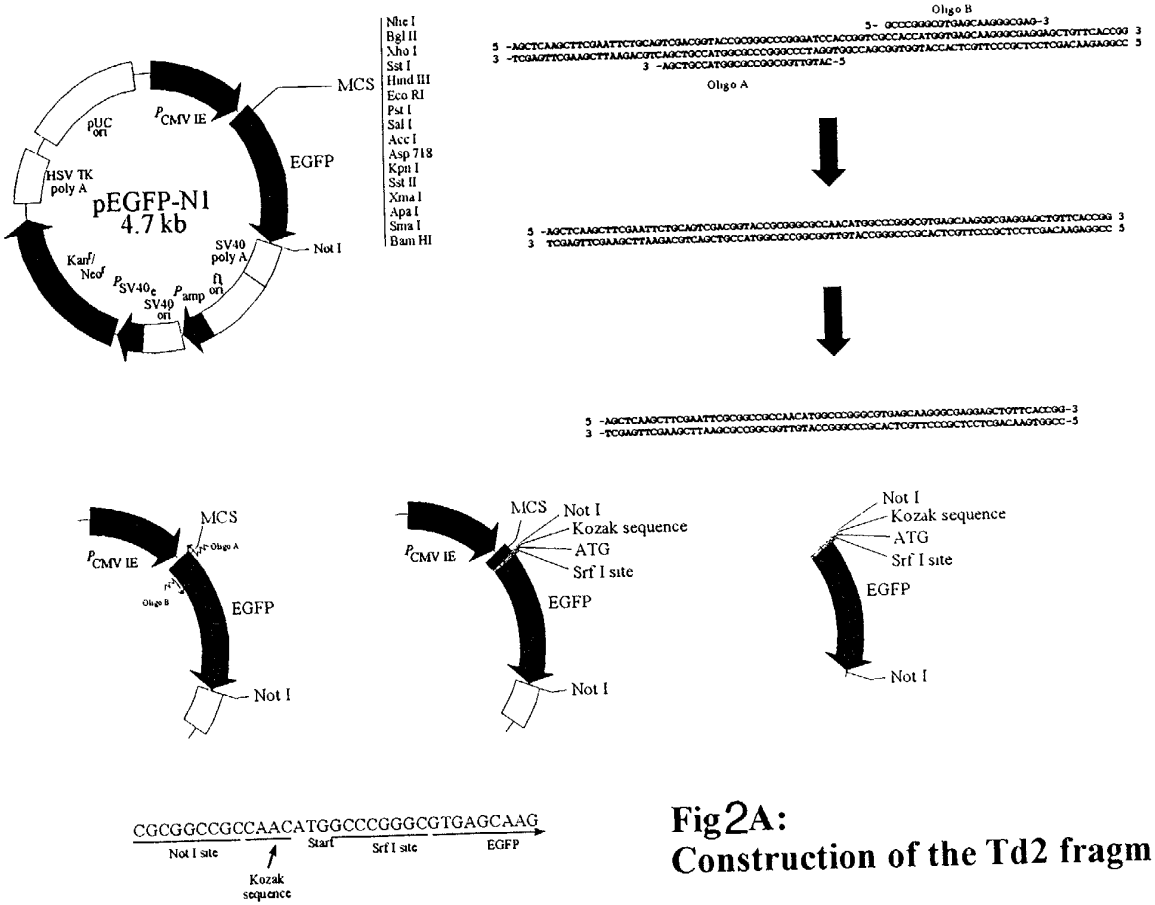


Fig2A:
Construction of the Td2 fragment

09/926201

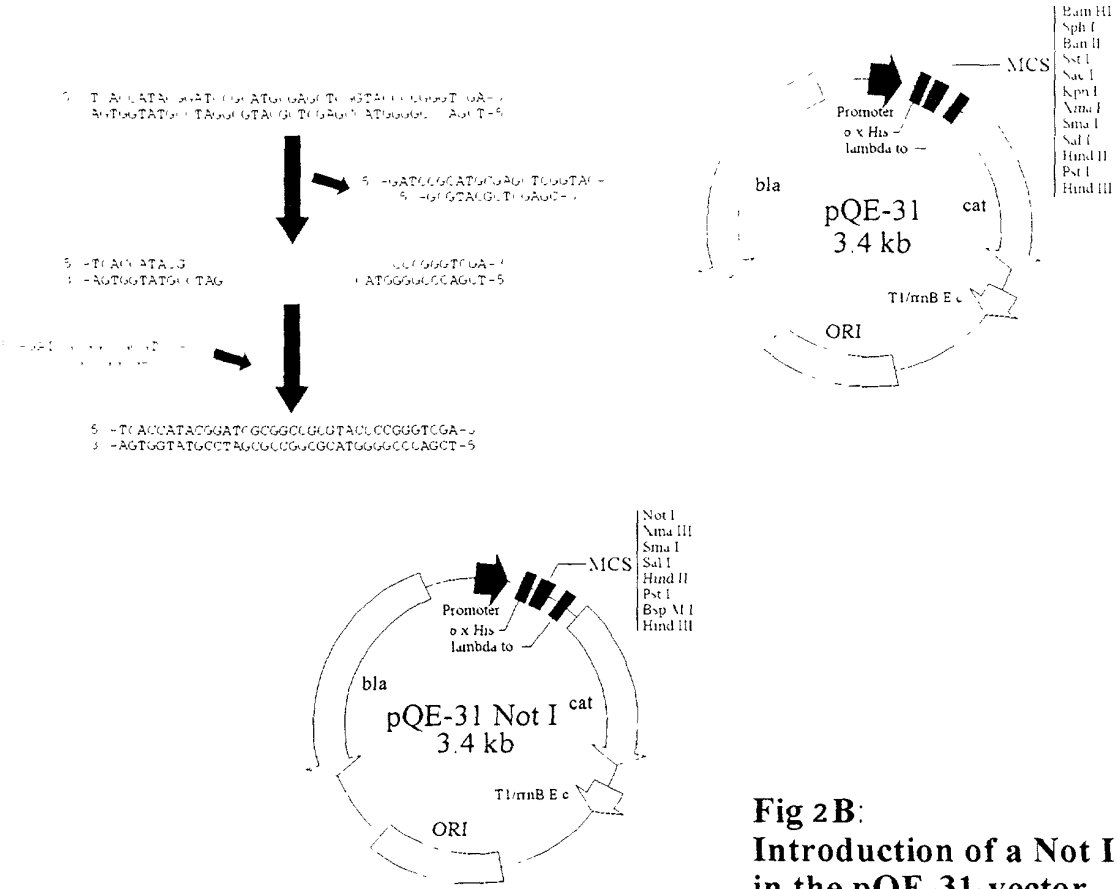


Fig 2B:
Introduction of a Not I site in the pQE-31-vector

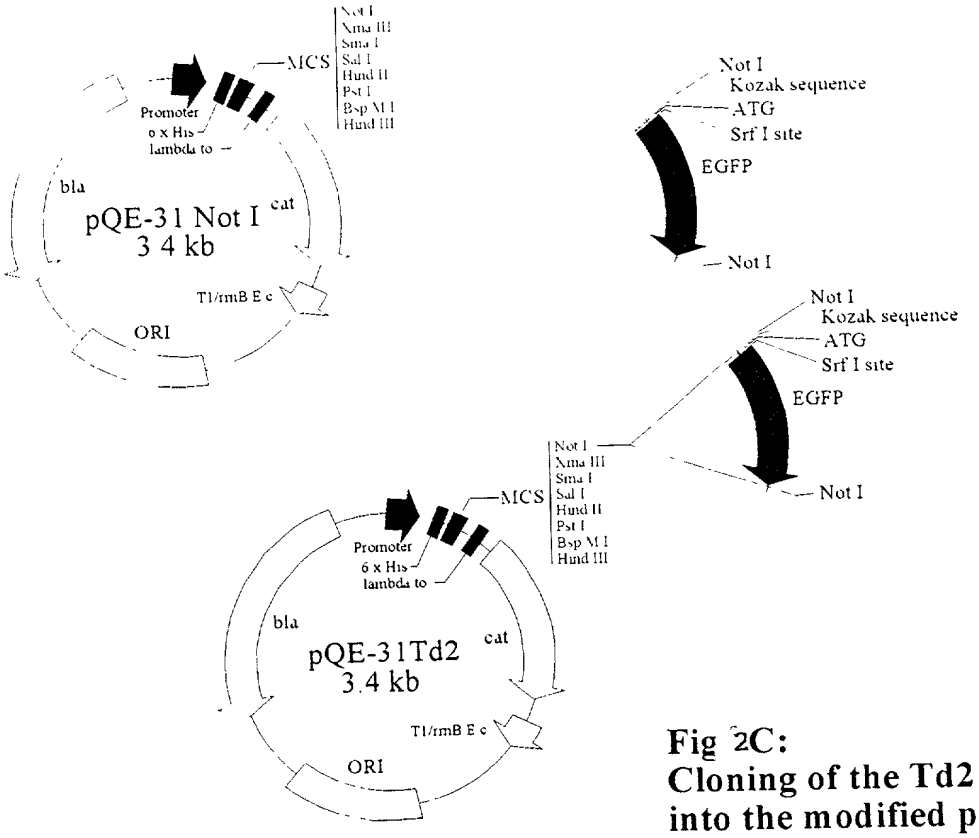


Fig 2C:
Cloning of the Td2 fragment into the modified pQE-31 vector

09/926201

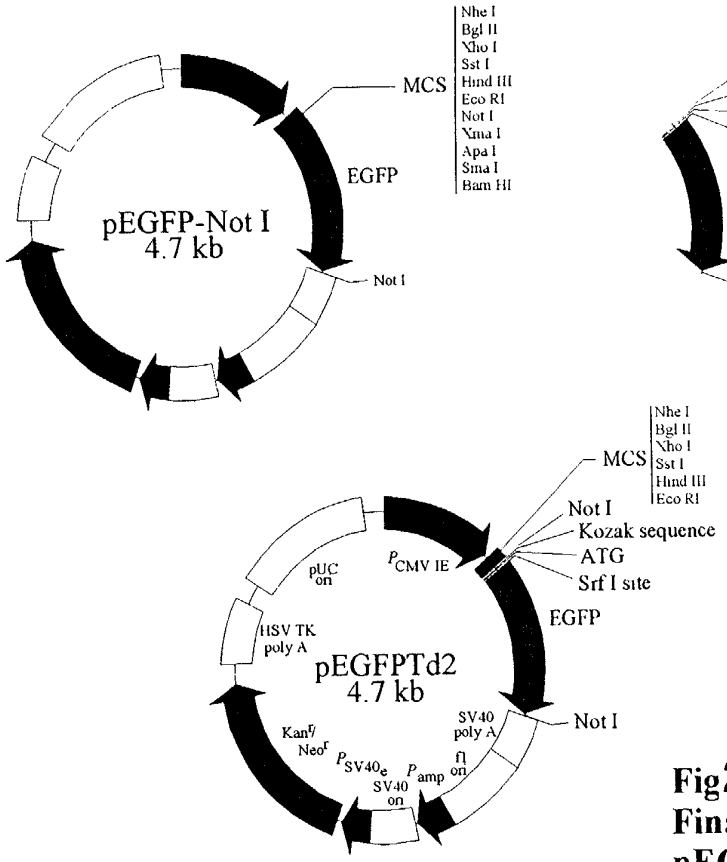


Fig2D:
Final step in the production of the pEGFPTd2 vector

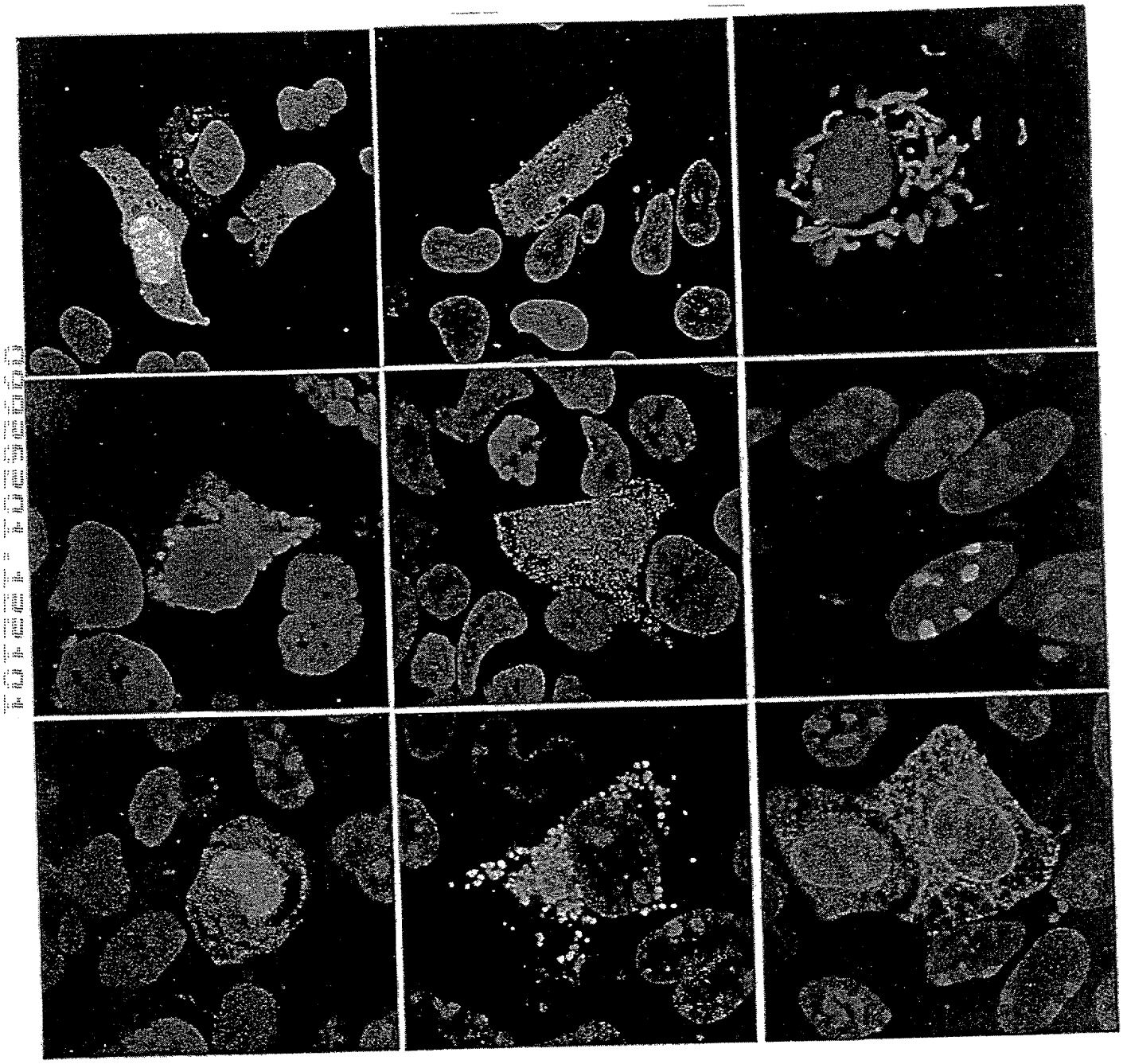


Figure 3

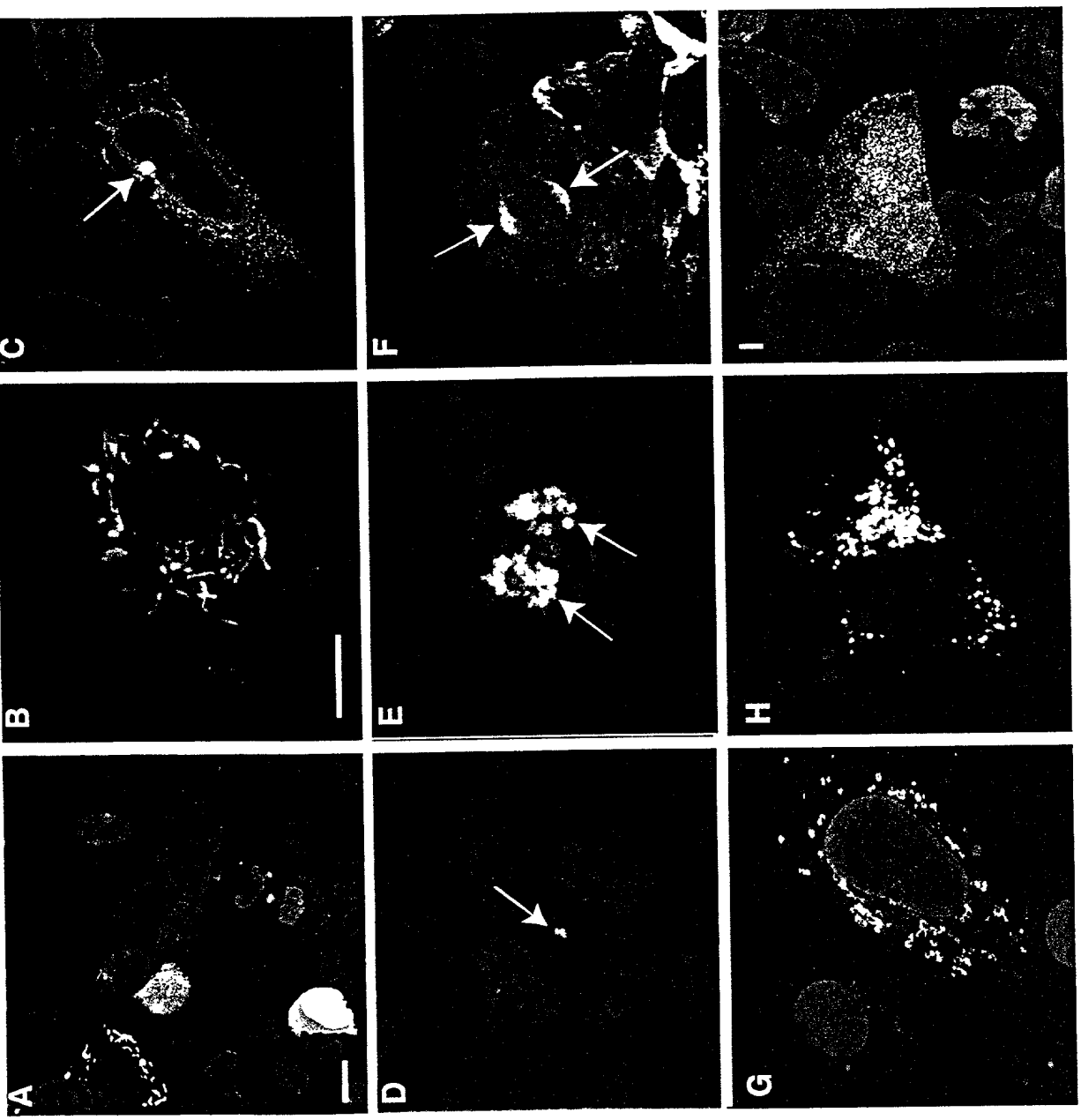
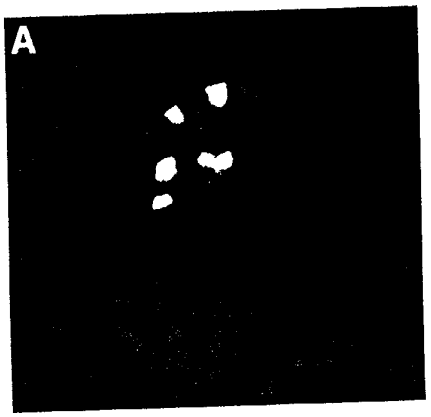
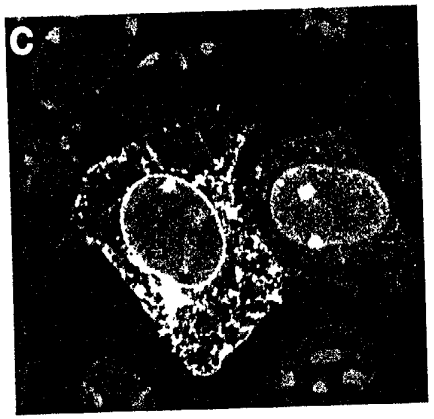


Figure 4



B

1 RKALYKRKYSAAKTKVEKKKKKEKV 25
26 LATVTKTVGGDKNGGTRVVKLRKMP 50
51 RYYPTEDVPRKLLSHGKKP 69



D

1 MAPMSIFIQLIYFLLFLFLGVICRC 25
26 SCIKENIRGV 35

Figure 5

09/926201

19741P US-WO
Docket No. 100564-00080

ARENT FOX KINTNER PLOTKIN & KAHN, PLLC
Nikaido, Marmelstein, Murray & Oram Intellectual Property Group

Declaration For U.S. Patent Application

As a below named inventor, I hereby declare that:
My residence, post office address and citizenship are as stated below my name.
I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled
(Insert Title) Detection, cloning and sequencing of polypeptides which drive the subcellular localization of proteins
the specification of which is attached hereto unless the following box is checked:

☒ was filed on March 23, 2000 as PCT International Application
Number PCT/EP 00/02607 and was amended on March 06, 2001
and/or was filed on _____ as United States Application
Number _____ and was amended on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claim(s), as amended by any amendment referred to above.
I acknowledge the duty to disclose information which is material to patentability as defined in 37 C.F.R. §1.56.
I hereby claim foreign priority benefits under 35 U.S.C. §119(a)-(d) or §365(b) of any foreign application(s) for patent or inventor's certificate, or §365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below any foreign application for patent or inventor's certificate or PCT International Application having a filing date before that of the application(s) for which priority is claimed:

(List prior foreign applications. See note A on back of this page)	(Number)	(Country)	(Day/Month/Year Filed)	Priority Claimed
	99 120 622.8	EP	Oct 18, 1999	☒ Yes ☐ No
	99 105 846.2	EP	March 23, 1999	☒ Yes ☐ No
				☐ Yes ☐ No

I hereby claim the benefit under 35 U.S.C. §119(c) of any United States provisional application(s) listed below.

(Application Number) (Filing Date)

(Application Number) (Filing Date)

(See Note B on back of this page) ☐ See attached list for additional prior foreign or provisional applications.

I hereby claim the benefit under 35 U.S.C. §120 of any United States application(s) or §365(c) of any PCT International application(s) designating the United States of America listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior application(s) (U.S. or PCT) in the manner provided by the first paragraph of 35, U.S.C. §112, I acknowledge the duty to disclose information which is material to patentability as defined in 37 C.F.R. §1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application.

(List prior U.S. Applications or PCT International applications designating the U.S.)	(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)
	(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)

And I hereby appoint as principal attorneys: Robert B. Murray, Reg. No. 22,980; Charles M. Marmelstein, Reg. No. 25,895; George E. Oram, Jr., Reg. No. 27,931; Douglas H. Goldhush, Reg. No. 33,125; David T. Nikaido, Reg. No. 22,663; Monica Chin Kitts, Reg. No. 36,105; Richard J. Berman, Reg. No. 39,107; King L. Wong, Reg. No. 37,500; James A. Poulos, III, Reg. No. 31,714; Patrick D. Muir, Reg. No. 37,403; Murat Ozgu, Reg. No. 44,275; Bradley D. Goldizen, Reg. No. 43,637, N. Alexander Nolte, Reg. No. 45,689 and Robert K. Carpenter, Reg. No. 34,794.

Please direct all communications to the following address:
ARENT FOX KINTNER PLOTKIN & KAHN, PLLC
1050 Connecticut Avenue, N.W., Suite 600
Washington, D.C. 20036-5339
Telephone No. (202) 857-5000; Facsimile No. (202) 638-4810

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

(See Note C on back of this page) 1-00

Full name of sole or first inventor GONZALEZ Cayetano
Inventor's signature X _____ Date X 21/3/01
Residence Erwin-Rohde Straße 22, 69120 Heidelberg, Germany DEX
Citizenship Spain
Post Office Address same as above

2-052 Full name of second joint inventor, if any BEJARANO Luis

Inventor's signature X Luis Bejarano X 21/09/01 Date

Residence Boxbergring 107, 69126 Heidelberg, Germany DEX

Citizenship Spain ✓

Post Office Address same as above

Full name of third joint inventor, if any _____

Inventor's signature _____ Date

Residence _____

Citizenship _____

Post Office Address _____

Full name of fourth joint inventor, if any _____

Inventor's signature _____ Date

Residence _____

Citizenship _____

Post Office Address _____

Full name of fifth joint inventor, if any _____

Inventor's signature _____ Date

Residence _____

Citizenship _____

Post Office Address _____

Full name of sixth joint inventor, if any _____

Inventor's signature _____ Date

Residence _____

Citizenship _____

Post Office Address _____

Full name of seventh joint inventor, if any _____

Inventor's signature _____ Date

Residence _____

Citizenship _____

Post Office Address _____

Full name of eighth joint inventor, if any _____

Inventor's signature _____ Date

Residence _____

Citizenship _____

Post Office Address _____

Full name of ninth joint inventor, if any _____

Inventor's signature _____ Date

Residence _____

Citizenship _____

Post Office Address _____

JC19 Rec'd PCT/PTO 24 SEP 2001
09/926201

SEQUENCE LISTING

<110> European Molecular Biology Laboratory (EMBL)

<120> Detection, cloning and sequencing of polypeptides which drive the subcellular localization of proteins

<130> 19741pwomd

<140> PCT/EP00/02607
<141> 2000-03-23

<150> EP 99105846.2
<151> 1999-03-23

<150> EP99120622.8
<151> 1999-10-18

<160> 8

<170> PatentIn Ver. 2.1

<210> 1
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: oligonucleotide (Oligo A)

<400> 1
catgttggcg gccgcggtac cgtcga 26

<210> 2
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: oligonucleotide (Oligo B)

<400> 2
gcccgggcgt gagcaagggc gag 23

<210> 3
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide (Oligo Not 1-1b)

<400> 3
gatcgcggcc gcgtac 16

<210> 4
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide (Oligo Not 1-8)

<400> 4
gcggccgc 8

<210> 5
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Fir

<400> 5
agcttcgaat tcgcggccgc caacatg 27

<210> 6
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Sec

<400> 6
tatgatctag agtcgcggcc gctttac 27

<210> 7
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Thi

<400> 7
tagcgctacc ggactcagat ctcgagc 27

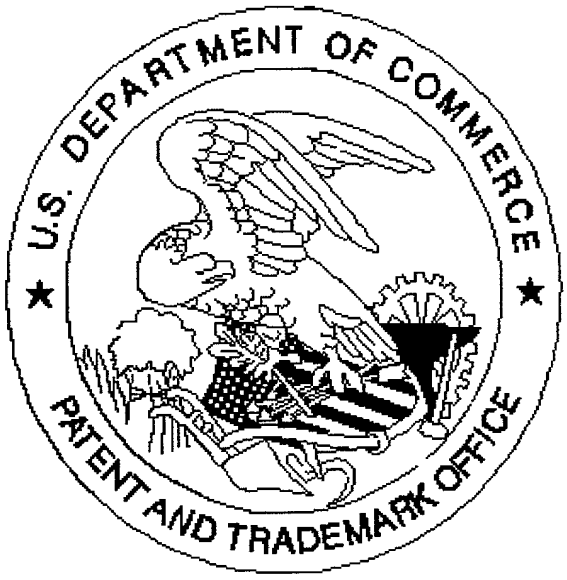
<210> 8
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Fou

<400> 8
aaaacctcta caaatgtggt atggctg 27

United States Patent & Trademark Office

Office of Initial Patent Examination -- Scanning Division



Application deficiencies found during scanning:

☐ Page(s)_____of_____ were not present for scanning. (Document title)

☐ Page(s)_____of_____were not present for scanning. (Document title)

☒ ***Scanned copy is best available.*** drawing figures are dark